



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jul 24, 2017 – 07:19 AM EDT

PDB ID : 5LMQ  
EMDB ID: : EMD-4076  
Title : Structure of bacterial 30S-IF1-IF3-mRNA-tRNA translation pre-initiation complex, open form (state-2A)  
Authors : Hussain, T.; Llacer, J.L.; Wimberly, B.T.; Ramakrishnan, V.  
Deposited on : unknown  
Resolution : 4.20 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

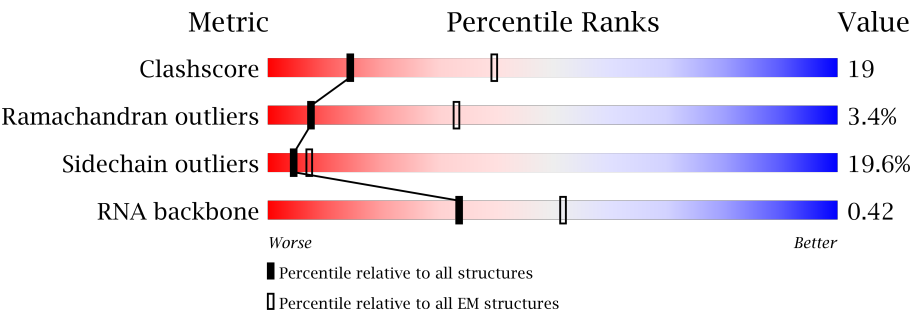
MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:  
*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.20 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.




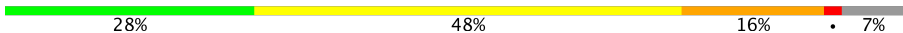



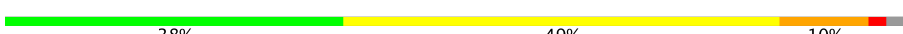











| Metric                | Whole archive (#Entries) | EM structures (#Entries) |
|-----------------------|--------------------------|--------------------------|
| Clashscore            | 125131                   | 1336                     |
| Ramachandran outliers | 121729                   | 1120                     |
| Sidechain outliers    | 121581                   | 1026                     |
| RNA backbone          | 3398                     | 335                      |

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$

| Mol | Chain | Length | Quality of chain   |
|-----|-------|--------|--|
| 1   | A     | 1522   | <div><div>26%</div><div>57%</div><div>16%</div><div>..</div></div> |
| 2   | B     | 256    | <div><div>45%</div><div>38%</div><div>7%</div><div>9%</div></div>  |
| 3   | C     | 239    | <div><div>50%</div><div>33%</div><div>14%</div></div>              |
| 4   | D     | 209    | <div><div>53%</div><div>39%</div><div>8%</div></div>               |
| 5   | E     | 162    | <div><div>52%</div><div>33%</div><div>7%</div><div>7%</div></div>  |
| 6   | F     | 101    | <div><div>58%</div><div>37%</div><div>..</div></div>               |
| 7   | G     | 156    | <div><div>65%</div><div>33%</div><div>..</div></div>               |
| 8   | H     | 138    | <div><div>65%</div><div>28%</div><div>7%</div></div>               |

Continued on next page...

Continued from previous page...

| Mol | Chain | Length | Quality of chain   |
|-----|-------|--------|--|
| 9   | I     | 128    |    |
| 10  | J     | 105    |    |
| 11  | K     | 129    |    |
| 12  | L     | 132    |    |
| 13  | M     | 126    |    |
| 14  | N     | 61     |    |
| 15  | O     | 89     |    |
| 16  | P     | 88     |    |
| 17  | Q     | 105    |    |
| 18  | R     | 88     |    |
| 19  | S     | 93     |    |
| 20  | T     | 106    |    |
| 21  | V     | 27     |  |
| 22  | W     | 72     |  |
| 23  | X     | 171    |  |
| 24  | Y     | 42     |  |
| 25  | Z     | 77     |  |

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol | Type | Chain | Res | Chirality | Geometry | Clashes | Electron density |
|-----|------|-------|-----|-----------|----------|---------|------------------|
| 27  | ZN   | D     | 300 | -         | -        | X       | -                |

## 2 Entry composition

There are 27 unique types of molecules in this entry. The entry contains 55841 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a RNA chain called 16S rRNA.

| Mol | Chain | Residues | Atoms |       |      |       |      | AltConf | Trace |
|-----|-------|----------|-------|-------|------|-------|------|---------|-------|
| 1   | A     | 1514     | Total | C     | N    | O     | P    | 0       | 0     |
|     |       |          | 32522 | 14481 | 6019 | 10512 | 1510 |         |       |

- Molecule 2 is a protein called 30S ribosomal protein S2.

| Mol | Chain | Residues | Atoms |      |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 2   | B     | 234      | Total | C    | N   | O   | S | 0       | 0     |
|     |       |          | 1900  | 1213 | 341 | 341 | 5 |         |       |

- Molecule 3 is a protein called 30S ribosomal protein S3.

| Mol | Chain | Residues | Atoms |      |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 3   | C     | 206      | Total | C    | N   | O   | S | 0       | 0     |
|     |       |          | 1612  | 1016 | 314 | 281 | 1 |         |       |

- Molecule 4 is a protein called 30S ribosomal protein S4.

| Mol | Chain | Residues | Atoms |      |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|-------|
| 4   | D     | 208      | Total | C    | N   | O   | S | 0       | 0     |
|     |       |          | 1703  | 1066 | 339 | 291 | 7 |         |       |

- Molecule 5 is a protein called 30S ribosomal protein S5.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 5   | E     | 150      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 1146  | 724 | 217 | 201 | 4 |         |       |

- Molecule 6 is a protein called 30S ribosomal protein S6.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 6   | F     | 101      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 843   | 531 | 155 | 154 | 3 |         |       |

- Molecule 7 is a protein called 30S ribosomal protein S7.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 7   | G     | 155      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 1257  | 781 | 252 | 218 | 6 |         |       |

- Molecule 8 is a protein called 30S ribosomal protein S8.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 8   | H     | 138      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 1116  | 705 | 215 | 193 | 3 |         |       |

- Molecule 9 is a protein called 30S ribosomal protein S9.

| Mol | Chain | Residues | Atoms |     |     |     | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---------|-------|
| 9   | I     | 127      | Total | C   | N   | O   | 0       | 0     |
|     |       |          | 1010  | 639 | 197 | 174 |         |       |

- Molecule 10 is a protein called 30S ribosomal protein S10.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 10  | J     | 98       | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 792   | 498 | 156 | 137 | 1 |         |       |

- Molecule 11 is a protein called 30S ribosomal protein S11.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 11  | K     | 122      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 906   | 563 | 172 | 168 | 3 |         |       |

- Molecule 12 is a protein called 30S ribosomal protein S12.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 12  | L     | 124      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 970   | 611 | 195 | 163 | 1 |         |       |

- Molecule 13 is a protein called 30S ribosomal protein S13.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 13  | M     | 125      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 997   | 617 | 207 | 171 | 2 |         |       |

- Molecule 14 is a protein called 30S ribosomal protein S14 type Z.

| Mol | Chain | Residues | Atoms |     |     |    |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|----|---|---------|-------|
| 14  | N     | 60       | Total | C   | N   | O  | S | 0       | 0     |
|     |       |          | 492   | 312 | 104 | 72 | 4 |         |       |

- Molecule 15 is a protein called 30S ribosomal protein S15.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 15  | O     | 88       | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 734   | 459 | 147 | 126 | 2 |         |       |

- Molecule 16 is a protein called 30S ribosomal protein S16.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 16  | P     | 83       | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 700   | 443 | 139 | 117 | 1 |         |       |

- Molecule 17 is a protein called 30S ribosomal protein S17.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 17  | Q     | 99       | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 823   | 528 | 151 | 142 | 2 |         |       |

- Molecule 18 is a protein called 30S ribosomal protein S18.

| Mol | Chain | Residues | Atoms |     |     |    | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|----|---------|-------|
| 18  | R     | 73       | Total | C   | N   | O  | 0       | 0     |
|     |       |          | 598   | 381 | 118 | 99 |         |       |

- Molecule 19 is a protein called 30S ribosomal protein S19.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 19  | S     | 82       | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 655   | 419 | 120 | 114 | 2 |         |       |

- Molecule 20 is a protein called 30S ribosomal protein S20.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 20  | T     | 99       | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 763   | 470 | 162 | 129 | 2 |         |       |

- Molecule 21 is a protein called 30S ribosomal protein Thx.

| Mol | Chain | Residues | Atoms |     |    |    | AltConf | Trace |
|-----|-------|----------|-------|-----|----|----|---------|-------|
| 21  | V     | 24       | Total | C   | N  | O  | 0       | 0     |
|     |       |          | 208   | 128 | 50 | 30 |         |       |

- Molecule 22 is a protein called Translation initiation factor IF-1.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 22  | W     | 71       | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 570   | 362 | 103 | 103 | 2 |         |       |

- Molecule 23 is a protein called Translation initiation factor IF-3.

| Mol | Chain | Residues | Atoms |     |     |     |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|-------|
| 23  | X     | 168      | Total | C   | N   | O   | S | 0       | 0     |
|     |       |          | 1356  | 853 | 249 | 245 | 9 |         |       |

- Molecule 24 is a RNA chain called mRNA.

| Mol | Chain | Residues | Atoms |     |    |     |    | AltConf | Trace |
|-----|-------|----------|-------|-----|----|-----|----|---------|-------|
| 24  | Y     | 21       | Total | C   | N  | O   | P  | 0       | 0     |
|     |       |          | 459   | 205 | 91 | 142 | 21 |         |       |

- Molecule 25 is a RNA chain called tRNA.

| Mol | Chain | Residues | Atoms |     |     |     |    |   | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|----|---|---------|-------|
| 25  | Z     | 77       | Total | C   | N   | O   | P  | S | 0       | 0     |
|     |       |          | 1643  | 735 | 297 | 534 | 76 | 1 |         |       |

- Molecule 26 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol | Chain | Residues | Atoms |    | AltConf |
|-----|-------|----------|-------|----|---------|
| 26  | W     | 1        | Total | Mg | 0       |
|     |       |          | 1     | 1  |         |
| 26  | A     | 63       | Total | Mg | 0       |
|     |       |          | 63    | 63 |         |

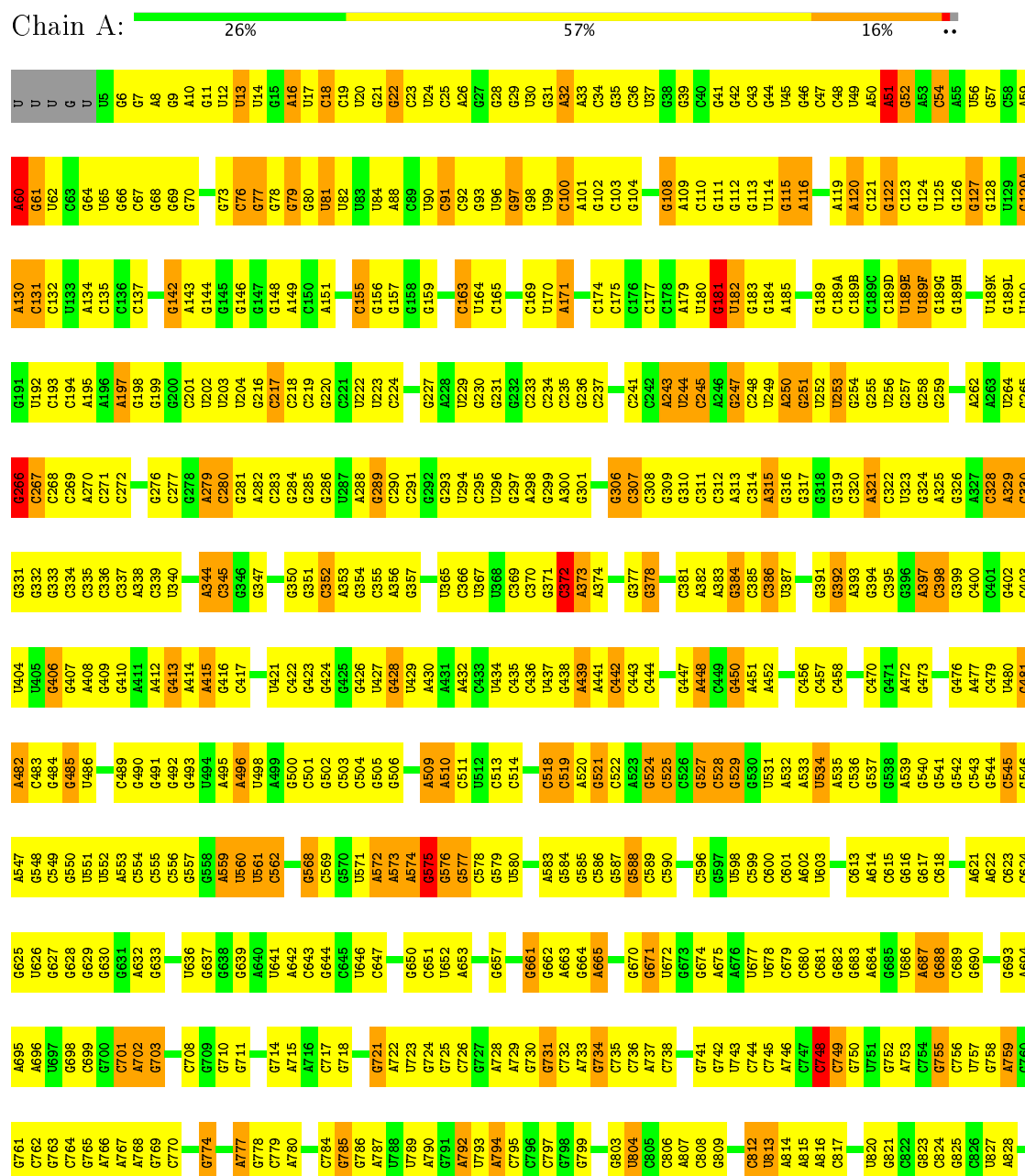
- Molecule 27 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol | Chain | Residues | Atoms |    | AltConf |
|-----|-------|----------|-------|----|---------|
| 27  | D     | 1        | Total | Zn | 0       |
|     |       |          | 1     | 1  |         |
| 27  | N     | 1        | Total | Zn | 0       |
|     |       |          | 1     | 1  |         |

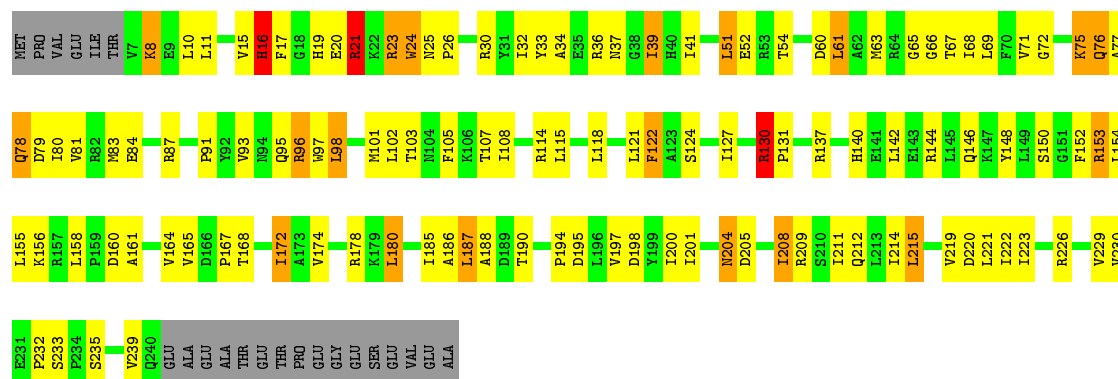
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

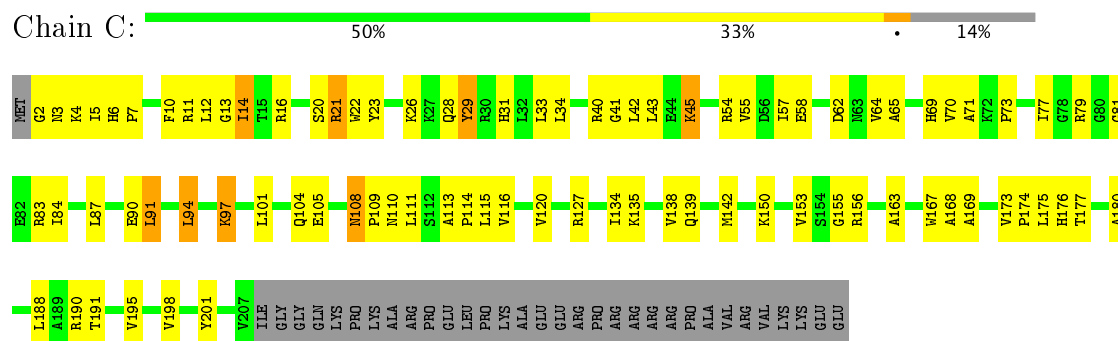
#### • Molecule 1: 16S rRNA



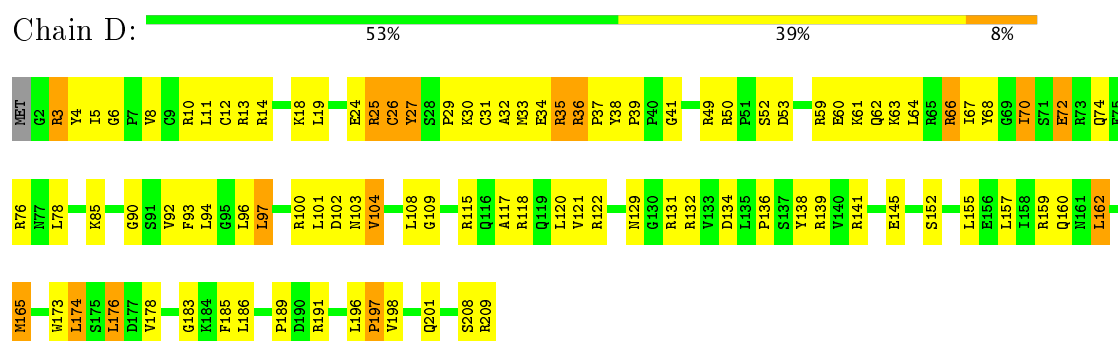




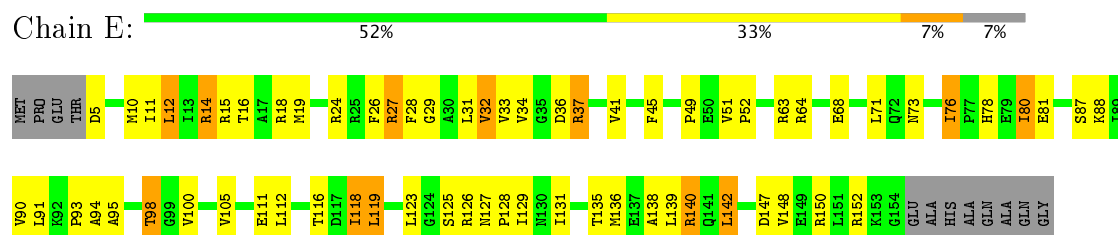
- Molecule 3: 30S ribosomal protein S3



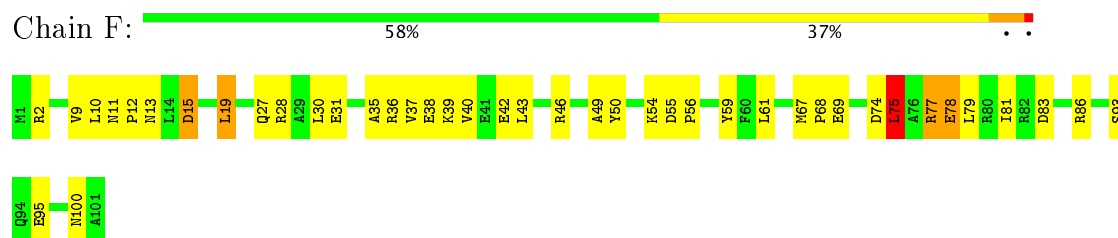
- Molecule 4: 30S ribosomal protein S4



- Molecule 5: 30S ribosomal protein S5

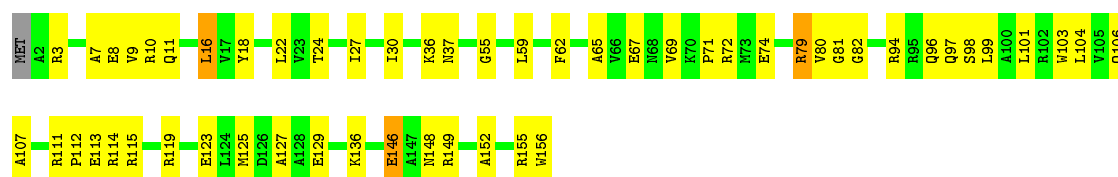


- Molecule 6: 30S ribosomal protein S6



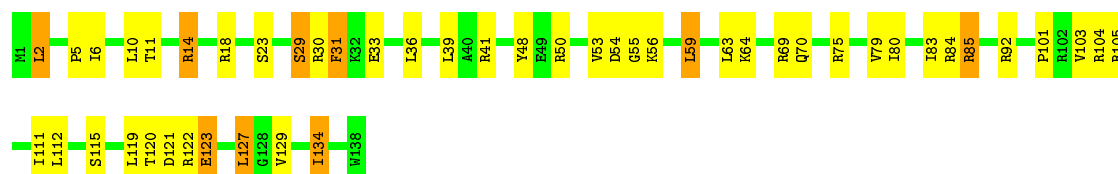
- Molecule 7: 30S ribosomal protein S7





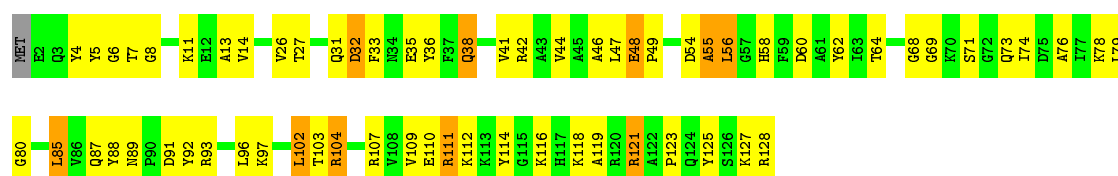
• Molecule 8: 30S ribosomal protein S8

Chain H: 65% 28% 7%



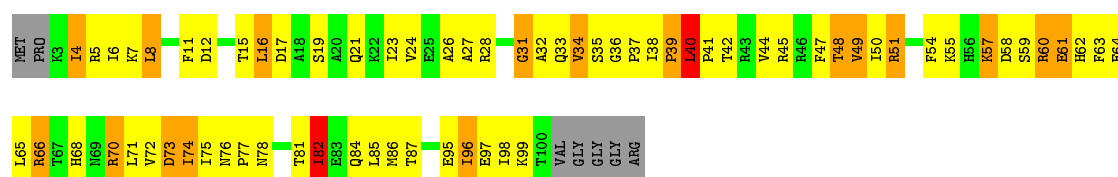
• Molecule 9: 30S ribosomal protein S9

Chain I: 48% 43% 8%



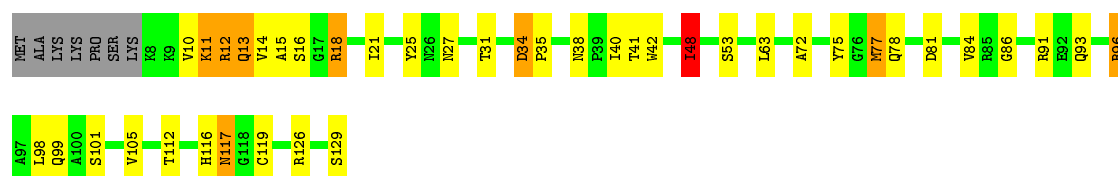
• Molecule 10: 30S ribosomal protein S10

Chain J: 28% 48% 16% 7%



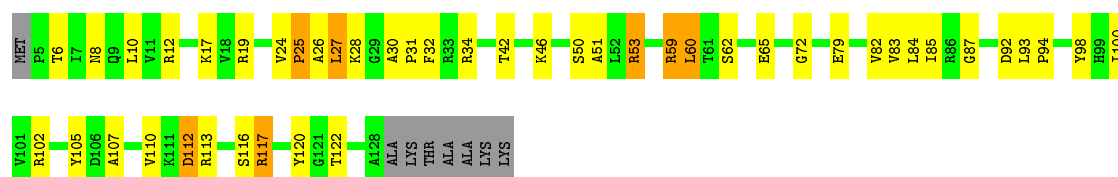
• Molecule 11: 30S ribosomal protein S11

Chain K: 63% 25% 6% 5%



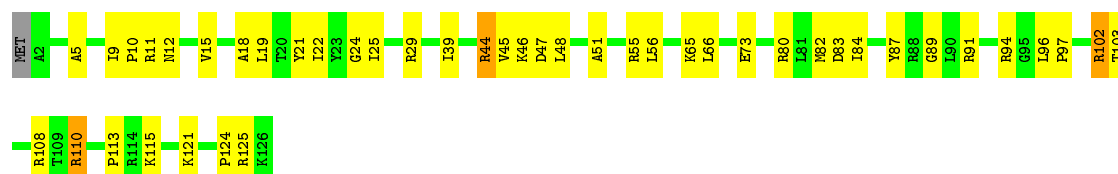
• Molecule 12: 30S ribosomal protein S12

Chain L: 59% 30% 5% 6%



- Molecule 13: 30S ribosomal protein S13

Chain M: 64% 33% ..



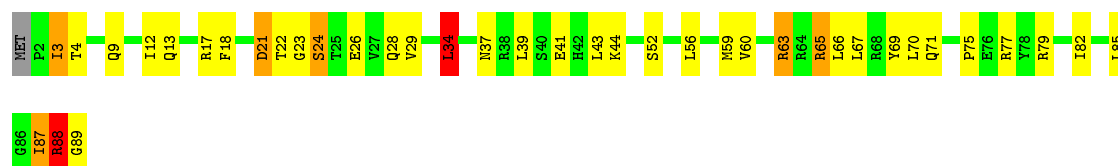
- Molecule 14: 30S ribosomal protein S14 type Z

Chain N: 38% 49% 10% ..



- Molecule 15: 30S ribosomal protein S15

Chain O: 55% 35% 7% ..



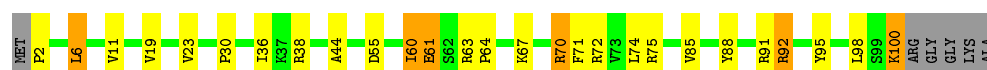
- Molecule 16: 30S ribosomal protein S16

Chain P: 57% 34% 6% ..



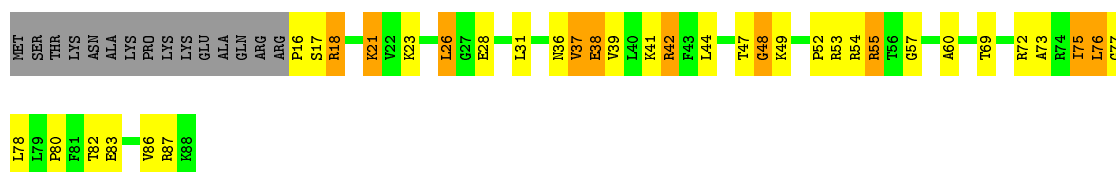
- Molecule 17: 30S ribosomal protein S17

Chain Q: 69% 20% 6% 6%



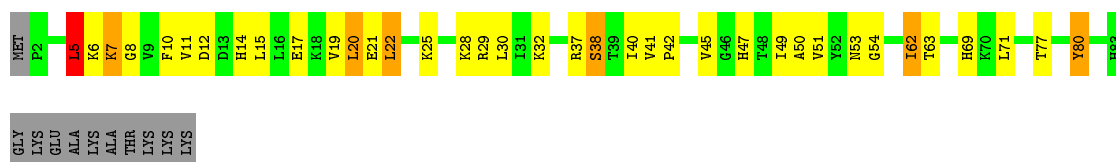
- Molecule 18: 30S ribosomal protein S18

Chain R: 42% 30% 11% 17%



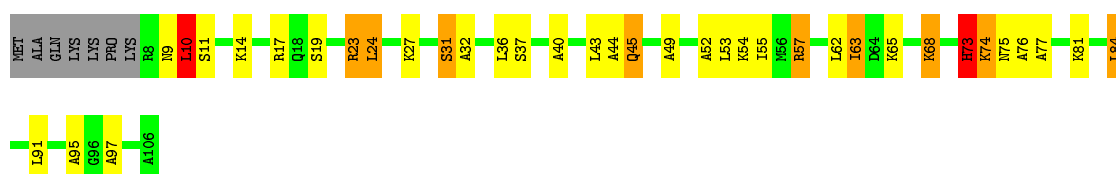
- Molecule 19: 30S ribosomal protein S19

Chain S:



- Molecule 20: 30S ribosomal protein S20

Chain T:



- Molecule 21: 30S ribosomal protein Thx

Chain V:



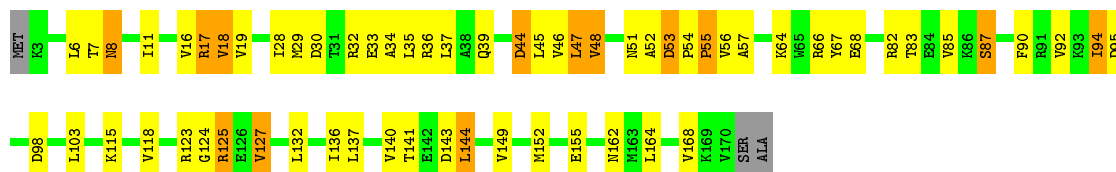
- Molecule 22: Translation initiation factor IF-1

Chain W:



- Molecule 23: Translation initiation factor IF-3

Chain X:



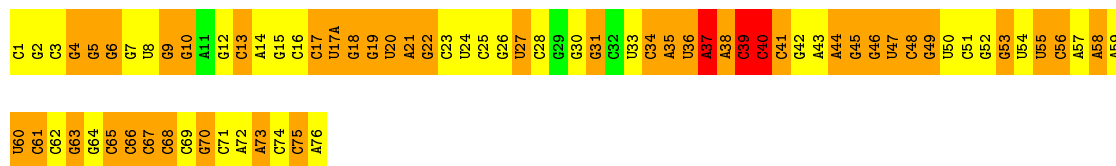
- Molecule 24: mRNA

Chain Y:  14% 21% 14% 50%



- Molecule 25: tRNA

Chain Z:  40% 52%



## 4 Experimental information

| Property                             | Value                                   | Source    |
|--------------------------------------|---|-----------|
| Reconstruction method                | SINGLE PARTICLE                         | Depositor |
| Imposed symmetry                     | POINT, Not provided                     | Depositor |
| Number of particles used             | 31888                                   | Depositor |
| Resolution determination method      | FSC 0.143 CUT-OFF                       | Depositor |
| CTF correction method                | PHASE FLIPPING AND AMPLITUDE CORRECTION | Depositor |
| Microscope                           | FEI POLARA 300                          | Depositor |
| Voltage (kV)                         | 300                                     | Depositor |
| Electron dose ( $e^-/\text{\AA}^2$ ) | 30                                      | Depositor |
| Minimum defocus (nm)                 | 1500                                    | Depositor |
| Maximum defocus (nm)                 | 3500                                    | Depositor |
| Magnification                        | 104478                                  | Depositor |
| Image detector                       | OTHER                                   | Depositor |

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 5MU, OMC, ZN, G7M, MG, 4SU, PSU

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Chain | Bond lengths |                | Bond angles |                 |
|-----|-------|--------------|----------------|-------------|-----------------|
|     |       | RMSZ         | # Z  >2        | RMSZ        | # Z  >2         |
| 1   | A     | 0.37         | 2/36394 (0.0%) | 0.74        | 24/56779 (0.0%) |
| 10  | J     | 0.51         | 0/805          | 0.76        | 0/1082          |
| 11  | K     | 0.49         | 0/921          | 0.82        | 0/1241          |
| 12  | L     | 0.44         | 0/986          | 0.77        | 0/1320          |
| 13  | M     | 0.53         | 0/1008         | 0.82        | 1/1347 (0.1%)   |
| 14  | N     | 0.44         | 0/501          | 0.80        | 1/664 (0.2%)    |
| 15  | O     | 0.52         | 1/745 (0.1%)   | 0.89        | 1/992 (0.1%)    |
| 16  | P     | 0.44         | 0/716          | 0.74        | 0/963           |
| 17  | Q     | 0.43         | 0/836          | 0.79        | 0/1117          |
| 18  | R     | 0.51         | 0/604          | 0.86        | 1/801 (0.1%)    |
| 19  | S     | 0.55         | 0/670          | 0.82        | 1/903 (0.1%)    |
| 2   | B     | 0.57         | 0/1935         | 0.86        | 1/2609 (0.0%)   |
| 20  | T     | 0.50         | 0/765          | 0.96        | 1/1007 (0.1%)   |
| 21  | V     | 0.47         | 0/212          | 0.77        | 0/277           |
| 22  | W     | 0.51         | 0/580          | 0.99        | 4/782 (0.5%)    |
| 23  | X     | 0.56         | 0/1375         | 0.86        | 1/1844 (0.1%)   |
| 24  | Y     | 0.56         | 0/516          | 0.83        | 0/804           |
| 25  | Z     | 0.55         | 0/1718         | 0.90        | 3/2678 (0.1%)   |
| 3   | C     | 0.48         | 0/1636         | 0.85        | 2/2205 (0.1%)   |
| 4   | D     | 0.46         | 0/1733         | 0.85        | 0/2318          |
| 5   | E     | 0.47         | 0/1162         | 0.82        | 0/1564          |
| 6   | F     | 0.47         | 0/856          | 0.88        | 1/1154 (0.1%)   |
| 7   | G     | 0.46         | 0/1276         | 0.79        | 0/1709          |
| 8   | H     | 0.43         | 0/1136         | 0.82        | 0/1527          |
| 9   | I     | 0.49         | 0/1029         | 0.79        | 1/1379 (0.1%)   |
| All | All   | 0.43         | 3/60115 (0.0%) | 0.78        | 43/89066 (0.0%) |

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.



| Mol | Chain | #Chirality outliers | #Planarity outliers |
|-----|-------|---------------------|---------------------|
| 12  | L     | 0                   | 1                   |
| 2   | B     | 0                   | 2                   |
| 23  | X     | 0                   | 2                   |
| 8   | H     | 0                   | 1                   |
| All | All   | 0                   | 6                   |

All (3) bond length outliers are listed below:

| Mol | Chain | Res  | Type | Atoms | Z     | Observed(Å) | Ideal(Å) |
|-----|-------|------|------|-------|-------|-------------|----------|
| 1   | A     | 999  | C    | O3'-P | -6.98 | 1.52        | 1.61     |
| 1   | A     | 1001 | A    | O3'-P | -6.75 | 1.53        | 1.61     |
| 15  | O     | 24   | SER  | CB-OG | 5.23  | 1.49        | 1.42     |

The worst 5 of 43 bond angle outliers are listed below:

| Mol | Chain | Res  | Type | Atoms       | Z      | Observed(°) | Ideal(°) |
|-----|-------|------|------|-------------|--------|-------------|----------|
| 22  | W     | 23   | ARG  | N-CA-C      | -10.10 | 83.74       | 111.00   |
| 22  | W     | 23   | ARG  | CB-CA-C     | -8.69  | 93.02       | 110.40   |
| 6   | F     | 75   | LEU  | CA-CB-CG    | 8.65   | 135.20      | 115.30   |
| 1   | A     | 1001 | A    | O4'-C4'-C3' | -8.10  | 95.90       | 104.00   |
| 1   | A     | 575  | G    | C2'-C3'-O3' | 7.99   | 127.07      | 109.50   |

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

| Mol | Chain | Res | Type | Group   |
|-----|-------|-----|------|---------|
| 2   | B     | 11  | LEU  | Peptide |
| 2   | B     | 130 | ARG  | Peptide |
| 8   | H     | 2   | LEU  | Peptide |
| 12  | L     | 112 | ASP  | Peptide |
| 23  | X     | 53  | ASP  | Peptide |

## 5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1   | A     | 32522 | 0        | 16435    | 1114    | 0            |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 2   | B     | 1900  | 0        | 1951     | 55      | 0            |
| 3   | C     | 1612  | 0        | 1677     | 81      | 0            |
| 4   | D     | 1703  | 0        | 1765     | 77      | 0            |
| 5   | E     | 1146  | 0        | 1207     | 74      | 0            |
| 6   | F     | 843   | 0        | 857      | 18      | 0            |
| 7   | G     | 1257  | 0        | 1296     | 25      | 0            |
| 8   | H     | 1116  | 0        | 1177     | 17      | 0            |
| 9   | I     | 1010  | 0        | 1037     | 47      | 0            |
| 10  | J     | 792   | 0        | 835      | 83      | 0            |
| 11  | K     | 906   | 0        | 928      | 21      | 0            |
| 12  | L     | 970   | 0        | 1057     | 23      | 0            |
| 13  | M     | 997   | 0        | 1072     | 27      | 0            |
| 14  | N     | 492   | 0        | 530      | 47      | 0            |
| 15  | O     | 734   | 0        | 771      | 20      | 0            |
| 16  | P     | 700   | 0        | 720      | 18      | 0            |
| 17  | Q     | 823   | 0        | 891      | 20      | 0            |
| 18  | R     | 598   | 0        | 670      | 25      | 0            |
| 19  | S     | 655   | 0        | 672      | 17      | 0            |
| 20  | T     | 763   | 0        | 861      | 20      | 0            |
| 21  | V     | 208   | 0        | 221      | 6       | 0            |
| 22  | W     | 570   | 0        | 599      | 24      | 0            |
| 23  | X     | 1356  | 0        | 1401     | 33      | 0            |
| 24  | Y     | 459   | 0        | 228      | 15      | 0            |
| 25  | Z     | 1643  | 0        | 844      | 128     | 0            |
| 26  | A     | 63    | 0        | 0        | 0       | 0            |
| 26  | W     | 1     | 0        | 0        | 0       | 0            |
| 27  | D     | 1     | 0        | 0        | 4       | 0            |
| 27  | N     | 1     | 0        | 0        | 0       | 0            |
| All | All   | 55841 | 0        | 39702    | 1833    | 0            |

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 19.

The worst 5 of 1833 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1         | Atom-2          | Interatomic distance (Å) | Clash overlap (Å) |
|----------------|-----------------|--------------------------|-------------------|
| 1:A:1080:A:H4' | 5:E:16:THR:CG2  | 1.32                     | 1.59              |
| 1:A:1080:A:C5' | 5:E:16:THR:HG21 | 1.42                     | 1.46              |
| 1:A:247:G:OP2  | 17:Q:100:LYS:CG | 1.64                     | 1.44              |
| 1:A:1080:A:C4' | 5:E:16:THR:CG2  | 2.00                     | 1.40              |
| 10:J:61:GLU:CD | 14:N:58:LYS:HD3 | 1.39                     | 1.39              |

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed        | Favoured   | Allowed   | Outliers | Percentiles |     |
|-----|-------|-----------------|------------|-----------|----------|-------------|-----|
| 2   | B     | 232/256 (91%)   | 187 (81%)  | 33 (14%)  | 12 (5%)  | 2           | 27  |
| 3   | C     | 204/239 (85%)   | 176 (86%)  | 20 (10%)  | 8 (4%)   | 3           | 33  |
| 4   | D     | 206/209 (99%)   | 185 (90%)  | 14 (7%)   | 7 (3%)   | 4           | 37  |
| 5   | E     | 148/162 (91%)   | 134 (90%)  | 12 (8%)   | 2 (1%)   | 13          | 54  |
| 6   | F     | 99/101 (98%)    | 90 (91%)   | 7 (7%)    | 2 (2%)   | 9           | 48  |
| 7   | G     | 153/156 (98%)   | 134 (88%)  | 15 (10%)  | 4 (3%)   | 6           | 42  |
| 8   | H     | 136/138 (99%)   | 122 (90%)  | 11 (8%)   | 3 (2%)   | 8           | 46  |
| 9   | I     | 125/128 (98%)   | 106 (85%)  | 16 (13%)  | 3 (2%)   | 7           | 44  |
| 10  | J     | 96/105 (91%)    | 74 (77%)   | 14 (15%)  | 8 (8%)   | 1           | 16  |
| 11  | K     | 120/129 (93%)   | 100 (83%)  | 14 (12%)  | 6 (5%)   | 2           | 28  |
| 12  | L     | 122/132 (92%)   | 99 (81%)   | 19 (16%)  | 4 (3%)   | 4           | 37  |
| 13  | M     | 123/126 (98%)   | 105 (85%)  | 15 (12%)  | 3 (2%)   | 7           | 44  |
| 14  | N     | 58/61 (95%)     | 49 (84%)   | 6 (10%)   | 3 (5%)   | 2           | 27  |
| 15  | O     | 86/89 (97%)     | 79 (92%)   | 6 (7%)    | 1 (1%)   | 15          | 57  |
| 16  | P     | 81/88 (92%)     | 73 (90%)   | 8 (10%)   | 0        | 100         | 100 |
| 17  | Q     | 97/105 (92%)    | 87 (90%)   | 9 (9%)    | 1 (1%)   | 18          | 61  |
| 18  | R     | 71/88 (81%)     | 64 (90%)   | 4 (6%)    | 3 (4%)   | 3           | 31  |
| 19  | S     | 80/93 (86%)     | 59 (74%)   | 17 (21%)  | 4 (5%)   | 2           | 28  |
| 20  | T     | 97/106 (92%)    | 84 (87%)   | 8 (8%)    | 5 (5%)   | 2           | 27  |
| 21  | V     | 22/27 (82%)     | 19 (86%)   | 3 (14%)   | 0        | 100         | 100 |
| 22  | W     | 69/72 (96%)     | 61 (88%)   | 6 (9%)    | 2 (3%)   | 5           | 40  |
| 23  | X     | 166/171 (97%)   | 143 (86%)  | 15 (9%)   | 8 (5%)   | 2           | 29  |
| All | All   | 2591/2781 (93%) | 2230 (86%) | 272 (10%) | 89 (3%)  | 7           | 37  |

5 of 89 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 2   | B     | 21  | ARG  |
| 2   | B     | 24  | TRP  |
| 4   | D     | 37  | PRO  |
| 9   | I     | 56  | LEU  |
| 13  | M     | 113 | PRO  |

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed       | Rotameric | Outliers | Percentiles |    |
|-----|-------|----------------|-----------|----------|-------------|----|
| 2   | B     | 202/220 (92%)  | 148 (73%) | 54 (27%) | 0           | 5  |
| 3   | C     | 160/188 (85%)  | 143 (89%) | 17 (11%) | 8           | 35 |
| 4   | D     | 180/181 (99%)  | 141 (78%) | 39 (22%) | 1           | 8  |
| 5   | E     | 115/123 (94%)  | 85 (74%)  | 30 (26%) | 0           | 5  |
| 6   | F     | 90/90 (100%)   | 71 (79%)  | 19 (21%) | 1           | 9  |
| 7   | G     | 126/127 (99%)  | 109 (86%) | 17 (14%) | 4           | 26 |
| 8   | H     | 119/119 (100%) | 90 (76%)  | 29 (24%) | 1           | 6  |
| 9   | I     | 98/99 (99%)    | 81 (83%)  | 17 (17%) | 2           | 16 |
| 10  | J     | 87/92 (95%)    | 65 (75%)  | 22 (25%) | 0           | 6  |
| 11  | K     | 92/99 (93%)    | 75 (82%)  | 17 (18%) | 2           | 14 |
| 12  | L     | 104/109 (95%)  | 85 (82%)  | 19 (18%) | 2           | 14 |
| 13  | M     | 100/101 (99%)  | 89 (89%)  | 11 (11%) | 7           | 33 |
| 14  | N     | 49/50 (98%)    | 37 (76%)  | 12 (24%) | 1           | 6  |
| 15  | O     | 79/80 (99%)    | 58 (73%)  | 21 (27%) | 0           | 5  |
| 16  | P     | 72/74 (97%)    | 56 (78%)  | 16 (22%) | 1           | 8  |
| 17  | Q     | 94/97 (97%)    | 84 (89%)  | 10 (11%) | 8           | 35 |
| 18  | R     | 64/77 (83%)    | 50 (78%)  | 14 (22%) | 1           | 8  |
| 19  | S     | 71/80 (89%)    | 56 (79%)  | 15 (21%) | 1           | 9  |
| 20  | T     | 76/82 (93%)    | 61 (80%)  | 15 (20%) | 1           | 12 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed        | Rotameric  | Outliers  | Percentiles |    |
|-----|-------|-----------------|------------|-----------|-------------|----|
| 21  | V     | 19/22 (86%)     | 16 (84%)   | 3 (16%)   | 3           | 21 |
| 22  | W     | 62/63 (98%)     | 51 (82%)   | 11 (18%)  | 2           | 16 |
| 23  | X     | 145/150 (97%)   | 122 (84%)  | 23 (16%)  | 3           | 20 |
| All | All   | 2204/2323 (95%) | 1773 (80%) | 431 (20%) | 4           | 12 |

5 of 431 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 8   | H     | 92  | ARG  |
| 10  | J     | 87  | THR  |
| 22  | W     | 19  | ASN  |
| 8   | H     | 121 | ASP  |
| 9   | I     | 109 | VAL  |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 6   | F     | 73  | ASN  |
| 7   | G     | 86  | GLN  |
| 19  | S     | 53  | ASN  |
| 7   | G     | 11  | GLN  |
| 3   | C     | 108 | ASN  |

### 5.3.3 RNA ⓘ

| Mol | Chain | Analysed        | Backbone Outliers | Pucker Outliers |
|-----|-------|-----------------|-------------------|-----------------|
| 1   | A     | 1506/1522 (98%) | 394 (26%)         | 107 (7%)        |
| 24  | Y     | 20/42 (47%)     | 8 (40%)           | 2 (10%)         |
| 25  | Z     | 76/77 (98%)     | 51 (67%)          | 9 (11%)         |
| All | All   | 1602/1641 (97%) | 453 (28%)         | 118 (7%)        |

5 of 453 RNA backbone outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1   | A     | 6   | G    |
| 1   | A     | 7   | G    |
| 1   | A     | 8   | A    |
| 1   | A     | 9   | G    |
| 1   | A     | 13  | U    |

5 of 118 RNA pucker outliers are listed below:

| Mol | Chain | Res  | Type |
|-----|-------|------|------|
| 1   | A     | 864  | A    |
| 1   | A     | 1032 | G    |
| 25  | Z     | 26   | G    |
| 1   | A     | 872  | A    |
| 1   | A     | 975  | A    |

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

5 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Type | Chain | Res | Link | Bond lengths |      |          | Bond angles |      |          |
|-----|------|-------|-----|------|--------------|------|----------|-------------|------|----------|
|     |      |       |     |      | Counts       | RMSZ | # Z  > 2 | Counts      | RMSZ | # Z  > 2 |
| 25  | OMC  | Z     | 32  | 25   | 15,22,23     | 0.88 | 1 (6%)   | 19,31,34    | 1.15 | 1 (5%)   |
| 25  | G7M  | Z     | 46  | 25   | 19,26,27     | 2.79 | 3 (15%)  | 19,39,42    | 2.44 | 6 (31%)  |
| 25  | 5MU  | Z     | 54  | 25   | 14,22,23     | 0.97 | 1 (7%)   | 16,32,35    | 3.10 | 4 (25%)  |
| 25  | PSU  | Z     | 55  | 25   | 16,21,22     | 0.85 | 0        | 20,30,33    | 3.70 | 8 (40%)  |
| 25  | 4SU  | Z     | 8   | 25   | 14,21,22     | 1.36 | 2 (14%)  | 15,30,33    | 2.08 | 5 (33%)  |

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

| Mol | Type | Chain | Res | Link | Chirals | Torsions  | Rings   |
|-----|------|-------|-----|------|---------|-----------|---------|
| 25  | OMC  | Z     | 32  | 25   | -       | 0/5/27/28 | 0/2/2/2 |
| 25  | G7M  | Z     | 46  | 25   | -       | 0/3/25/26 | 0/3/3/3 |
| 25  | 5MU  | Z     | 54  | 25   | -       | 0/3/25/26 | 0/2/2/2 |
| 25  | PSU  | Z     | 55  | 25   | -       | 0/7/25/26 | 0/2/2/2 |
| 25  | 4SU  | Z     | 8   | 25   | -       | 0/3/25/26 | 0/2/2/2 |

The worst 5 of 7 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms   | Z     | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|---------|-------|-------------|----------|
| 25  | Z     | 8   | 4SU  | C4-S4   | -3.58 | 1.60        | 1.67     |
| 25  | Z     | 8   | 4SU  | O4'-C1' | 2.23  | 1.44        | 1.41     |
| 25  | Z     | 32  | OMC  | O4'-C1' | 2.43  | 1.44        | 1.41     |
| 25  | Z     | 54  | 5MU  | O4'-C1' | 2.78  | 1.45        | 1.41     |
| 25  | Z     | 46  | G7M  | C6-C5   | 4.43  | 1.49        | 1.41     |

The worst 5 of 24 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms    | Z     | Observed(°) | Ideal(°) |
|-----|-------|-----|------|----------|-------|-------------|----------|
| 25  | Z     | 55  | PSU  | N1-C2-N3 | -9.79 | 121.36      | 128.40   |
| 25  | Z     | 55  | PSU  | C5-C4-N3 | -8.50 | 118.45      | 125.43   |
| 25  | Z     | 54  | 5MU  | C5-C4-N3 | -6.39 | 118.20      | 125.24   |
| 25  | Z     | 46  | G7M  | C5-C6-N1 | -5.65 | 115.44      | 123.48   |
| 25  | Z     | 8   | 4SU  | C5-C4-N3 | -3.78 | 118.95      | 123.73   |

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 8 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 25  | Z     | 46  | G7M  | 4       | 0            |
| 25  | Z     | 54  | 5MU  | 2       | 0            |
| 25  | Z     | 55  | PSU  | 1       | 0            |
| 25  | Z     | 8   | 4SU  | 1       | 0            |

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 66 ligands modelled in this entry, 66 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

| Mol | Chain | Number of breaks |
|-----|-------|------------------|
| 1   | A     | 6                |

The worst 5 of 6 chain breaks are listed below:

| Model | Chain | Residue-1 | Atom-1 | Residue-2 | Atom-2 | Distance (Å) |
|-------|-------|-----------|--------|-----------|--------|--------------|
| 1     | A     | 84:U      | O3'    | 88:A      | P      | 4.82         |
| 1     | A     | 93:G      | O3'    | 96:U      | P      | 4.82         |
| 1     | A     | 841:U     | O3'    | 848:C     | P      | 4.25         |
| 1     | A     | 204:U     | O3'    | 216:G     | P      | 3.81         |
| 1     | A     | 1442(A):G | O3'    | 1442(B):A | P      | 3.32         |